An $\varepsilon-$regularity result for optimal transport maps between continuous densities

M. Goldman *

November 9, 2020

Abstract

The aim of this short note is to extend the recent variational proof of partial regularity for optimal transport maps to the case of continuous densities.

1 Introduction

The aim of this short note is to extend the partial regularity result for optimal transport maps obtained in [5] to the case of continuous densities (rather than Hölder continuous). The interest lies in the proof rather than in the result in itself since it is known to hold under the weaker assumption that the densities are bounded from above and below [2]. Indeed, we show that for the squared Euclidean cost, both the variational approach to regularity theory for the Monge-Ampère equation recently developed in [5, 4, 6] and the one of [1] lead to the same result. We must however emphasize that the major achievement of [1] is the treatment of arbitrary cost functions (see [7] for the extension of the variational approach to that setting). Our main $\varepsilon-$regularity theorem is the following (compare with [1, Th. 4.3]):

Theorem 1.1. *Let ρ_0 and ρ_1 be densities with compact support*[i] *and T be the optimal transport map from ρ_0 to ρ_1 for the squared Euclidean cost on $\mathbb{R}^d$. For every $\alpha \in (0,1)$, there exists $\varepsilon(\alpha, d) > 0$ such that if for some $R > 0$,*

$$\frac{1}{(2R)^{d+2}} \int_{B_{2R}} |T - x|^2 \rho_0 + \|1 - \rho_0\|^2_{L^\infty(B_{2R})} + \|1 - \rho_1\|^2_{L^\infty(B_{2R})} \le \varepsilon,$$

then T is of class $C^{0,\alpha}$ in B_R.

*Université de Paris, CNRS, Sorbonne-Université, Laboratoire Jacques-Louis Lions (LJLL), F-75005 Paris, France, michael.goldman@u-paris.fr

[i] we assume compactness of the supports for simplicity. The statement is valid as soon as an optimal transport map exists (see [9]).

With this ε-regularity result at hand and arguing as for [5, Th. 1.1], it is not hard to prove that T is a $C^{0,\alpha}$ homeomorphism outside of a set of measure zero if ρ_0 and ρ_1 are continuous.

Theorem 1.2. *For E and F two bounded open sets, let $\rho_0 : E \to \mathbb{R}^+$ and $\rho_1 : F \to \mathbb{R}^+$ be two continuous densities with equal masses, both bounded and bounded away from zero and let T be the optimal transport map between ρ_0 and ρ_1. Then, there exist open sets $E' \subseteq E$ and $F' \subseteq F$ with $|E\backslash E'| = |F\backslash F'| = 0$ and such that for every $\alpha \in (0,1)$, T is a $C^{0,\alpha}$ homeomorphism between E' and F'.*

The proof of Theorem 1.1 follows very closely the proof of [5, Th. 1.2]. It is based on a Campanato iteration scheme which uses at his heart an harmonic approximation result (see Proposition 3.1). The main difference with [5, Th. 1.2] lies in the iteration argument (see Theorem 3.5 below). Indeed, for continuous densities the linear part of the affine transformations introduced in the excess improvement by tilting estimate do not necessarily converge to the identity. This causes the possible blow-up of the $C^{1,\alpha}-$norms.

2 Notation

In the paper we will use the following notation. The symbols $\sim$, $\gtrsim$, $\lesssim$ indicate estimates that hold up to a global constant C, which typically only depends on the dimension d and the Hölder exponent α (if applicable). For instance, $f \lesssim g$ means that there exists such a constant with $f \le Cg$, $f \sim g$ means $f \lesssim g$ and $g \lesssim f$. An assumption of the form $f \ll 1$ means that there exists $\varepsilon > 0$, typically only depending on dimension and the Hölder exponent, such that if $f \le \varepsilon$, then the conclusion holds. We write $|E|$ for the Lebesgue measure of a set E. When no confusion is possible, we will drop the integration measures in the integrals. For $R > 0$ and $x_0 \in \mathbb{R}^d$, $B_R(x_0)$ denotes the ball of radius R centered in x_0. When $x_0 = 0$, we will simply write B_R for $B_R(0)$. We will also use the notation

$$⨍_{B_R} f := \frac{1}{|B_R|} \int_{B_R} f.$$

Let ρ_0 and ρ_1 be two densities with compact support in $\mathbb{R}^d$ and equal mass. We say that T is an optimal transport map between ρ_0 and ρ_1 if it minimizes

$$W^2(\rho_0, \rho_1) := \min_{T\sharp\rho_0=\rho_1} \int_{\mathbb{R}^d} |T - x|^2 \rho_0, \tag{2.1}$$

where by a slight abuse of notation $T\sharp\rho_0$ denotes the push-forward by T of the measure $\rho_0 dx$. We refer the reader to [9] for the existence, uniqueness and characterization of such maps.

3 Proof of Theorem 1.1

Let T be the minimizer of (2.1). As in [5], the proof of Theorem 1.1 is based on the decay properties of the excess energy

$$\mathcal{E}(\rho_0, \rho_1, T, R) := R^{-2} \fint_{B_R} |T - x|^2 \rho_0. \tag{3.1}$$

As already alluded to, the main ingredient for the proof of Theorem 1.1 is the following harmonic approximation result (which by scaling we state for $R = 1$).

Proposition 3.1. *For every $0 < \tau \ll 1$, there exists $\varepsilon(\tau, d) > 0$ and $C(\tau, d) > 0$ such that if*

$$\mathcal{E}(\rho_0, \rho_1, T, 1) + \|1 - \rho_0\|^2_{L^\infty(B_1)} + \|1 - \rho_1\|^2_{L^\infty(B_1)} \leq \varepsilon, \tag{3.2}$$

then there exists a function φ harmonic in $B_{1/2}$, such that

$$\int_{B_{1/2}} |T - (x + \nabla\varphi)|^2 \rho_0 \leq \tau \mathcal{E}(\rho_0, \rho_1, T, 1) + C \left(\|1 - \rho_0\|^2_{L^\infty(B_1)} + \|1 - \rho_1\|^2_{L^\infty(B_1)} \right) \tag{3.3}$$

and

$$\sup_{B_{1/2}} |\nabla\varphi|^2 \lesssim \mathcal{E}(\rho_0, \rho_1, T, 1) + \|1 - \rho_0\|^2_{L^\infty(B_1)} + \|1 - \rho_1\|^2_{L^\infty(B_1)}. \tag{3.4}$$

Proof. To simplify notation let $\mathcal{E} := \mathcal{E}(\rho_0, \rho_1, T, 1)$ and $D := \|1-\rho_0\|^2_{L^\infty(B_1)} + \|1-\rho_1\|^2_{L^\infty(B_1)}$. The claim is an almost direct application of [4, Th. 1.5]. Notice first that for $i = 0, 1$,

$$W^2 \left(\rho_i \llcorner B_1, \frac{\rho_i(B_1)}{|B_1|} \chi_{B_1} \right) + \left(\frac{\rho_i(B_1)}{|B_1|} - 1 \right)^2 \lesssim \|1 - \rho_i\|^2_{L^\infty(B_1)}. \tag{3.5}$$

Therefore, [4, Th. 1.5] gives the existence of a radius $R \in (3/4, 4/5)$, a constant $c \in \mathbb{R}$ and a couple (ρ, j) solving in the distributional sense the continuity equation[ii]

$$\partial_t \rho + \nabla \cdot j = 0 \ \text{ on } \mathbb{R}^d \times (0,1) \qquad \text{and} \qquad \rho(\cdot, 0) = \rho_0, \ \rho(\cdot, 1) = \rho_1 \tag{3.6}$$

such that the following holds. If Φ solves

$$\Delta\Phi = c \ \text{ in } B_R \qquad \text{and} \qquad \nu \cdot \nabla\Phi = \nu \cdot \int_0^1 j dt \ \text{ on } \partial B_R,$$

where ν denotes the external normal to ∂B_R, then

$$\int_{B_{1/2}} |T - (x + \nabla\Phi)|^2 \rho_0 \leq \tau\mathcal{E} + CD$$

[ii]note that (ρ, j) is actually the solution of the Eulerian version of (2.1), see [4].

and

$$\sup_{B_{1/2}} |\nabla\Phi|^2 \lesssim \mathcal{E} + D.$$

Using (3.6) and integration by parts, it is readily seen that we must have

$$c = \fint_{B_R} (\rho_0 - \rho_1)$$

and thus $|c|^2 \lesssim D$. Taking $\varphi := \Phi - \frac{c}{2d}|x|^2$ and using triangle inequality we get (3.3) and (3.4). ☐

Remark 3.2. *Instead of appealing to [4, Th. 1.5] whose proof is quite long and intricate, one could alternatively give a direct proof of Proposition 3.1 following almost verbatim the proof of [5, Prop. 3.5]. One would only need to replace the use of [5, Lem. 2.2], which required the densities to be Hölder continuous, by Lemma 3.3 below. With respect to* (3.3)*, this would lead to the slightly more quantitative statement*

$$\int_{B_{1/2}} |T - (x + \nabla\varphi)|^2 \rho_0 \lesssim \mathcal{E}(\rho_0, \rho_1, T, 1)^{\frac{d+2}{d+1}} + \|1 - \rho_0\|^2_{L^\infty(B_1)} + \|1 - \rho_1\|^2_{L^\infty(B_1)}.$$

Lemma 3.3. *For $g \in L^\infty(B_1)$, every solution solution φ of*

$$\Delta\varphi = g \quad \text{in } B_1 \qquad \text{and} \qquad \nu \cdot \nabla\varphi = \frac{1}{\mathcal{H}^{d-1}(\partial B_1)} \int_{B_1} g \quad \text{on } \partial B_1,$$

satisfies

$$\sup_{B_1} |\nabla\varphi|^2 \lesssim \|g\|^2_{L^\infty(B_1)}.$$

Proof. This follows from global Schauder estimates [8, Th. 3.16 (iii)] and the fact that if $g \in L^\infty(B_1)$, then g is in the Morrey space $L^{2,d-2(1-\alpha)}(B_1)$ for every $0 < \alpha < 1$ (see [8]). ☐

We now prove that as in [5, Prop. 3.6], this estimate implies an "excess improvement by tilting"-estimate. Even though the proof is similar to the one in [5], we include it for the reader's convenience.

Proposition 3.4. *For every $\beta \in (0, 1)$ there exist $\varepsilon(d, \beta) > 0$, $\theta = \theta(d, \beta) > 0$ and $C_\theta(d, \beta) > 0$ with the property that for every $R > 0$ such that*

$$\mathcal{E}(\rho_0, \rho_1, T, R) + \|1 - \rho_0\|^2_{L^\infty(B_R)} + \|1 - \rho_1\|^2_{L^\infty(B_R)} \le \varepsilon, \tag{3.7}$$

there exist a symmetric matrix M with $\det M = 1$ *and a vector b with*

$$|M - Id|^2 + \frac{1}{R^2}|b|^2 \lesssim \mathcal{E}(\rho_0, \rho_1, T, R) + \|1 - \rho_0\|^2_{L^\infty(B_R)} + \|1 - \rho_1\|^2_{L^\infty(B_R)}, \tag{3.8}$$

such that, letting $\hat{x} := M^{-1}x$, $\hat{y} := M(y-b)$ and then

$$\hat{T}(\hat{x}) := M(T(x)-b), \quad \hat{\rho}_0(\hat{x}) := \rho_0(x) \quad \text{and} \quad \hat{\rho}_1(\hat{y}) := \rho_1(y), \tag{3.9}$$

we have

$$\mathcal{E}(\hat{\rho}_0, \hat{\rho}_1, \hat{T}, \theta R) \leq \theta^{2\beta}\mathcal{E}(\rho_0, \rho_1, T, R) + C_\theta \left(\|1-\rho_0\|^2_{L^\infty(B_R)} + \|1-\rho_1\|^2_{L^\infty(B_R)} \right). \tag{3.10}$$

Proof. By a rescaling $\widetilde{x} = R^{-1}x$, which amounts to the re-definition $\widetilde{T}(\widetilde{x}) := R^{-1}T(R\widetilde{x})$ (which preserves optimality) and $\widetilde{b} := R^{-1}b$, we may assume that $R = 1$.
As above, we introduce the notation

$$\mathcal{E} := \mathcal{E}(\rho_0, \rho_1, T, 1) \qquad \text{and} \qquad D := \|1-\rho_0\|^2_{L^\infty(B_1)} + \|1-\rho_1\|^2_{L^\infty(B_1)}.$$

Let $\tau \in (0,1)$ to be fixed later and then φ be the harmonic function given by Proposition 3.1. Define $b := \nabla\varphi(0)$, $A := \nabla^2\varphi(0)$ and set $M := e^{-A/2}$, so that $\det M = 1$. Using (3.4) from Proposition 3.1 and the mean value property for harmonic functions, we see that (3.8) is satisfied.

Defining $\hat{\rho}_i$ and $\hat{T}$ as in (3.9) we have by (3.8) and (3.7)

$$\begin{aligned}
\fint_{B_\theta} |\hat{T} - \hat{x}|^2 \hat{\rho}_0 &= \fint_{MB_\theta} |M(T-b) - M^{-1}x|^2 \rho_0 \\
&\lesssim \fint_{B_{2\theta}} |T - (M^{-2}x + b)|^2 \rho_0 \\
&\lesssim \fint_{B_{2\theta}} |T - (x + \nabla\varphi)|^2 \rho_0 + \fint_{B_{2\theta}} |(M^{-2} - Id - A)x|^2 \rho_0 \\
&\qquad + \fint_{B_{2\theta}} |\nabla\varphi - b - Ax|^2 \rho_0 \\
&\lesssim \fint_{B_{2\theta}} |T - (x + \nabla\varphi)|^2 \rho_0 + \theta^2 |M^{-2} - Id - A|^2 + \sup_{B_{2\theta}} |\nabla\varphi - b - Ax|^2.
\end{aligned}$$

Recalling $M = e^{-A/2}$, $A = \nabla^2\varphi(0)$, and $b = \nabla\varphi(0)$, we obtain

$$\begin{aligned}
\theta^{-2}\fint_{B_\theta} |\hat{T} - x|^2 \hat{\rho}_0 &\overset{(3.3)}{\lesssim} \theta^{-(d+2)} \left(\tau\mathcal{E} + C_\tau D\right) + |\nabla^2\varphi(0)|^4 + \theta^2 \sup_{B_{2\theta}} |\nabla^3\varphi|^2 \\
&\overset{(3.4)}{\lesssim} \theta^{-(d+2)} \left(\tau\mathcal{E} + C_\tau D\right) + (\mathcal{E} + D)^2 + \theta^2 (\mathcal{E} + D) \\
&\lesssim \left(\tau\theta^{-(d+2)} + \theta^2\right)\mathcal{E} + C_\tau \theta^{-(d+2)} D,
\end{aligned}$$

where we used the harmonicity of $\nabla\varphi$ and the fact that $\mathcal{E} + D \ll \theta^2$ (recall that θ has not been fixed yet). We may thus find a constant $C(d) > 0$ such that

$$\theta^{-2}\fint_{B_\theta} |\hat{T} - x|^2 \hat{\rho}_0 \leq C\left(\tau\theta^{-(d+2)} + \theta^2\right)\mathcal{E} + C_\tau \theta^{-(d+2)} D.$$

We now fix $\theta(d,\beta)$ such that $C\theta^2 \leq \frac{1}{2}\theta^{2\beta}$, which is possible because $\beta < 1$. We finally choose $\tau \ll 1$ such that also $C\tau\theta^{-(d+2)} \leq \frac{1}{2}\theta^{2\beta}$, which concludes the proof of (3.10). □

We may finally prove our main $\varepsilon-$regularity result. As already pointed out in the introduction, it is in this iteration argument that the proof departs from the one in [5]. Indeed, under the assumption that the densities are merely continuous, the distance to the identity of the linear transformations M in (3.8) are not decaying and we need to compensate the possible blow-up of the cumulated linear transformations by downgrading the $C^{1,\alpha}$ estimates to $C^{0,\alpha}$ estimates. A similar argument is used in [1]. Notice that the Campanato iteration itself is somewhat simpler here compared to [5, Prop. 3.7] since we do not need to introduce an extra dilation factor at every step to propagate the smallness assumption on the data.

Theorem 3.5. *For every $\alpha \in (0,1)$, if*

$$\mathcal{E}(\rho_0, \rho_1, T, 2R) + \|1-\rho_0\|^2_{L^\infty(B_{2R})} + \|1-\rho_1\|^2_{L^\infty(B_{2R})} \ll 1, \tag{3.11}$$

then T is of class $C^{0,\alpha}$ in B_R.

Proof. By scale invariance, we may assume that $R = 1$. Let us fix $\alpha \in (0,1)$. By Campanato's theory, see [3, Th. 5.5], we have to prove that (3.11) implies

$$\sup_{x_0\in B_1} \sup_{r\le \frac{1}{2}} \min_b \frac{1}{r^{2\alpha}} \fint_{B_r(x_0)} |T-b|^2 \lesssim 1. \tag{3.12}$$

Let us first notice that (3.11) implies that for every $x_0 \in B_1$

$$\mathcal{E} := \fint_{B_1(x_0)} |T-x|^2\rho_0 \ll 1 \qquad \text{and} \qquad \|1-\rho_0\|^2_{L^\infty(B_1(x_0))} + \|1-\rho_1\|^2_{L^\infty(B_1(x_0))} \ll 1. \tag{3.13}$$

Therefore, in order to prove (3.12), it is enough to show that (3.13) implies that for $r \le \frac{1}{2}$,

$$\min_b \fint_{B_r(x_0)} |T-b|^2 \lesssim r^{2\alpha}. \tag{3.14}$$

Without loss of generality we may now assume that $x_0 = 0$. To simplify notation, we let

$$\varepsilon := \mathcal{E} + \|1-\rho_0\|^2_{L^\infty(B_1)} + \|1-\rho_1\|^2_{L^\infty(B_1)}.$$

Fix from now on $\beta \in (0,1)$ and let $\theta(d,\beta)$ be given by Proposition 3.4. Thanks to (3.13), Proposition 3.4 applies and there exist a (symmetric) matrix M_1 of unit determinant and a vector b_1 such that $T_1(x) := B_1(T(M_1x)-b_1)$, $\rho_0^1(x) := \rho_0(M_1x)$ and $\rho_1^1(x) := \rho_1(M_1^{-1}x+b_1)$ satisfy

$$\mathcal{E}_1 := \mathcal{E}(\rho_0^1, \rho_1^1, T_1, \theta) \le \theta^{2\beta}\mathcal{E} + C_\theta \left(\|1-\rho_0\|^2_{L^\infty(B_1)} + \|1-\rho_1\|^2_{L^\infty(B_1)}\right) \le (\theta^{2\beta} + C_\theta)\varepsilon. \tag{3.15}$$

If T is a minimizer of (2.1), then so is T_1 with (ρ_0, ρ_1) replaced by (ρ_0^1, ρ_1^1). Indeed, because $\det M_1 = 1$, T_1 sends ρ_0^1 on ρ_1^1 and if T is the gradient of a convex function ψ then $T_1 = \nabla\psi_1$ where $\psi_1(x) := \psi(M_1x) - b_1 \cdot M_1x$ is also a convex function, which characterizes optimality

[9, Th. 2.12]. Moreover, since by (3.8), we have $|M_1 - Id|^2 \lesssim \varepsilon$ and $|b_1|^2 \lesssim \theta^2\varepsilon$, if ε is small enough then $M_1 B_\theta \subseteq B_1$ and $M_1^{-1}B_\theta + b_1 \subseteq B_1$,

$$\|1-\rho_0^1\|^2_{L^\infty(B_\theta)} + \|1-\rho_1^1\|^2_{L^\infty(B_\theta)} \le \|1-\rho_0\|^2_{L^\infty(B_1)} + \|1-\rho_1\|^2_{L^\infty(B_1)} \le \varepsilon. \tag{3.16}$$

Therefore, we may iterate Proposition 3.4, $K > 1$ times to find a sequence of (symmetric) matrices M_k with $\det M_k = 1$, a sequence of vectors b_k and a sequence of maps T_k such that setting for $1 \le k \le K$,

$$T_k(x) := M_k(T_{k-1}(M_k x) - b_k), \quad \rho_0^k(x) := \rho_0^{k-1}(M_k x) \quad \text{and} \quad \rho_1^k(x) := \rho_1^{k-1}(M_k^{-1}x + b_k),$$

we have

$$\|1-\rho_0^k\|^2_{L^\infty(B_{\theta^k})} + \|1-\rho_1^k\|^2_{L^\infty(B_{\theta^k})} \le \varepsilon \tag{3.17}$$

$$\mathcal{E}_k := \mathcal{E}(\rho_0^k, \rho_1^k, T_k, \theta^k) \le \theta^{2\beta}\mathcal{E}_{k-1} + C_\theta \varepsilon, \tag{3.18}$$

$$|M_k - Id|^2 \lesssim \mathcal{E}_{k-1} + \varepsilon, \tag{3.19}$$

$$\frac{1}{\theta^{2(k-1)}}|b_k|^2 \lesssim \mathcal{E}_{k-1} + \varepsilon. \tag{3.20}$$

A simple induction argument shows that from (3.18) we get

$$\mathcal{E}_k \le \theta^{2k\beta}\mathcal{E} + C_\theta \sum_{j=0}^{k-1} \theta^{2\beta j}\varepsilon \lesssim \varepsilon \tag{3.21}$$

and so (3.19) and (3.20) lead to

$$\max(|M_k|^2, |M_k^{-1}|^2) \le (1 + C\sqrt{\varepsilon}) \qquad \text{and} \qquad |b_k|^2 \lesssim \theta^{2k}\varepsilon. \tag{3.22}$$

In particular, this implies that $M_K B_{\theta^K} \subseteq B_{\theta^{K-1}}$ and $M_K^{-1}B_{\theta^K} + b_K \subseteq B_{\theta^{K-1}}$ so that we may keep iterating Proposition 3.4.

Letting, $A_k := M_k M_{k-1} \cdots M_1$ and $d_k := \sum_{i=1}^k M_k M_{k-1} \cdots M_i b_i$, we see that $T_k(x) = A_k T(A_k^* x) - d_k$. By (3.22),

$$\max(|A_k|^2, |A_k^{-1}|^2) \le (1 + C\sqrt{\varepsilon})^k. \tag{3.23}$$

We first estimate by definition of T_k, the fact that $\det A_k = 1$ and $\|1-\rho_0^k\|_{L^\infty(B_{\theta^k})} \ll 1$,

$$\begin{aligned}
\fint_{A_k^*(B_{\theta^k})} |T + A_k^{-1}d_k|^2 &\lesssim \fint_{A_k^*(B_{\theta^k})} |T - A_k^{-1}A_k^{-*}x + A_k^{-1}d_k|^2 + \fint_{A_k^*(B_{\theta^k})} |A_k^{-1}A_k^{-*}x|^2 \\
&\lesssim \fint_{B_{\theta^k}} |A_k^{-1}(T_k - x)|^2 + \theta^{2k}|A_k^{-1}|^2 \\
&\lesssim |A_k^{-1}|^2\,(\mathcal{E}_k + 1)\,\theta^{2k} \\
&\overset{(3.23)\&(3.21)}{\lesssim} (1 + C\sqrt{\varepsilon})^k \theta^{2k}.
\end{aligned}$$

Now if ε is small enough, (3.23) yields $B_{\frac{1}{2}\left(\frac{\theta}{1+C\sqrt{\varepsilon}}\right)^k} \subseteq A_k^*(B_{\theta^k})$ and therefore

$$\min_b \fint_{B_{\frac{1}{2}\left(\frac{\theta}{1+C\sqrt{\varepsilon}}\right)^k}} |T-b|^2 \lesssim (1+C\sqrt{\varepsilon})^{kd} \fint_{A_k^*(B_{\theta^k})} |T + A_k^{-1} d_k|^2$$
$$\lesssim (1+C\sqrt{\varepsilon})^{k(d+2)} \theta^{2k}.$$

Since α and θ are fixed, if ε is small enough, then $1 + C\sqrt{\varepsilon} \leq \theta^{-\frac{2(1-\alpha)}{d+2(1+\alpha)}}$ so that

$$\theta^2 (1+C\sqrt{\varepsilon})^{d+2} \leq \left(\frac{\theta}{1+C\sqrt{\varepsilon}}\right)^{2\alpha}$$

From this (3.14) follows, which concludes the proof of (3.12).

□

Acknowledgment

I warmly thank G. De Philippis for suggesting that our proof from [5] could cover the case of merely continuous densities. I also thank A. Zilio for pointing out the reference [8]. Part of this research was supported by the project ANR-18-CE40-0013 SHAPO financed by the French Agence Nationale de la Recherche (ANR) and by the LYSM LIA AMU CNRS ECM INdAM. I also thank the Centro de Giorgi in Pisa for its hospitality.

References

[1] G. De Philippis and A. Figalli, *Partial regularity for optimal transport maps*, Publ. Math. Inst. Hautes Études Sci. **121** (2015), 81–112.

[2] A. Figalli and Y.-H. Kim, *Partial regularity of Brenier solutions of the Monge-Ampère equation*, Discrete Contin. Dyn. Syst. **28** (2010), no. 2, 559–565.

[3] M. Giaquinta and L. Martinazzi, *An introduction to the regularity theory for elliptic systems, harmonic maps and minimal graphs*, second ed., Appunti. Scuola Normale Superiore di Pisa (Nuova Serie), vol. 11, Edizioni della Normale, Pisa, 2012.

[4] M. Goldman, M. Huesmann, and F. Otto, *Quantitative linearization results for the Monge-Ampère equation*, arXiv e-prints (2019), arXiv:1905.09678.

[5] M. Goldman and F. Otto, *A variational proof of partial regularity for optimal transportation maps*, Annales de l'ENS (to appear).

[6] T. Miura and F. Otto, *Sharp boundary ε-regularity of optimal transport maps*, arXiv e-prints (2020), arXiv:2002.08668.

[7] F. Otto, M. Prod'homme, and T. Ried, *Variational approach to regularity of optimal transport maps: general cost functions*, arXiv e-prints (2020), arXiv:2007.11549.

[8] G. M. Troianiello, *Elliptic differential equations and obstacle problems*, The University Series in Mathematics, Plenum Press, New York, 1987.

[9] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003.